

Formale Grundlagen der Sozioökonomie

Termin 7

Jakob Kapeller

University of Duisburg-Essen
Institute for Socio-Economics &

Johannes Kepler University Linz
Institute for Comprehensive Analysis of the Economy (ICAE)

Editor: *Heterodox Economics Newsletter*

www.jakob-kapeller.org | www.uni-due.de | www.heterodoxnews.com



Open-Minded



Motivation

“Refusing to deal with numbers rarely serves the interest of the least well-off.”

Piketty, 2014

Agenda

- Wachstumsraten und Verwendung des Logarithmus
- Rechnen mit Erwartungswerten
- Differentialrechnung, insb. Optimierung
- Lineare Algebra, insb. Matrizenalgebra

Wachstumsraten und Logarithmus

Berechnung von Wachstumsraten

$$\underbrace{x_t}_{\text{x... Wachstumsrate von X}} = \frac{X_t - X_{t-1}}{X_{t-1}} = \frac{X_t}{\underbrace{X_{t-1}}_{\text{t... Zeitindex}}} - 1$$

- **Vorsicht:** „relativer Anteil“ vs. „Prozent (%)“ vs. „Prozentpunkte“
- z.B: X wächst von 100 auf 120, d.h. $(120/100) - 1 = 0.2$
- 0.2 = Anteil des Zuwachses relativ zur Periode (t-1) entspricht 20% Wachstum
- Nächstes Jahr wächst X auf 138, d.h. $(138/120) - 1 = 0.15$ (15% Wachstum)
- Also ist die Wachstumsrate x um „5 Prozentpunkte“ gefallen.
- **Was passiert bei mehreren Perioden?**
- $X_1 = X_0 (1 + x)$
- $X_3 = X_2 (1 + x) = X_0 (1 + x) (1 + x) (1 + x) = X_0 (1 + x)^3$
- $X_t = X_0 (1 + x)^t$

Wachstumsraten mit mehreren Einflussfaktoren

- Beispiel: Produzierte Gütermenge ($Y^A = Y^A/H * H$)
 - Hängt ab von: Zahl der Arbeitsstunden H und deren Produktivität $PR (= Y^A/H)$
 - Beschäftigte wachsen um 10%, Produktivität ebenso um 10%, Arbeitszeit/Beschäftigtem bleibt konstant — Was passiert mit Produktion?
- Näherungsformel allgemein:
 - $(1 + a) * (1 + b) = 1 + a + b + ab \approx 1 + a + b$
 - wenn die Wachstumsraten a, b eher klein sind, dann können diese addiert werden!
- Näherungsformel Beispiel Produktion:
 - $(1 + 0.1) * (1 + 0.1) = 1 + 0.1 + 0.1 + 0.01 = 1.21 \approx 1.2$
- Alternatives Beispiel: Wirtschaftswachstum und Inflation
 - $(1 + y_{nom}) = (1 + y_{real}) \cdot (1 + \pi) = 1 + y_{real} + \pi + y_{real} \cdot \pi \approx 1 + y_{real} + \pi$

$$\begin{array}{ccc} +? \% & +10 \% & +10 \% \\ \uparrow & \uparrow & \uparrow \\ Y^A & = & Y^A/H * H \end{array}$$

Arithmetisches und geometrisches Mittel

$$\bar{z}_a = \frac{1}{n} \sum_{i=1}^n z_i = \frac{1}{n} (z_1 + z_2 + \dots + z_n)$$

$$\bar{z}_g = \sqrt[n]{\prod_{i=1}^n z_i} = \sqrt[n]{z_1 \cdot z_2 \cdot \dots \cdot z_n}$$

$z = (1 + x)$

- Periodisches Wachstum: Multiplikativer Prozess → **Geometrisches Mittel**
 - Beispiel:** 100 € Investment mit 50% Wachstum in Jahr 1 (= 150 €) und 50% Verlust im Jahr 2 (= 75 €). Arithmetisches Mittel = 0, aber das reale Wachstum ist -25%!
 - Für geometrisches Mittel: Umformen von $X_t = X_0(1 + x)^t$ auf $\left(\frac{X_t}{X_0}\right)^{1/t} = \sqrt[t]{\frac{X_t}{X_0}} = (1 + x) = \bar{z}_g$
 - Äquivalent zur Definition des geometrischen Mittels, da $X_t = X_0(1 + x_1)(1 + x_2) \dots (1 + x_t)$
 - Anwendung auf unser Beispiel: $(75\text{€} / 100\text{€})^{1/2} - 1 = (-13.3 \%)$ bzw. $(1.5 \cdot 0.5)^{1/2} = 0.866$
- Erweiterung: Diskretes vs. Kontinuierliches Wachstum**

$$X_t = X_0(1 + x)^t \approx X_0 e^{xt} \Leftrightarrow \frac{X_t}{X_0} \approx e^{xt} \Leftrightarrow \frac{\ln(X_t) - \ln(X_0)}{t} \approx x$$

Differenz zweier Logarithmen:
Ergibt (näherungsweise) die
relative Änderung.

Änderungsraten und Verwendung des Log's

- Also: die **logarithmische Differenz** gibt die **relative Änderung** an

$$x_t = \frac{X_t - X_{t-1}}{X_{t-1}} = \frac{X_t}{X_{t-1}} - 1 \approx \ln X_t - \ln X_{t-1}$$

- Das ergibt sich aus den vorangegangenen Folien – Alternative Herleitung: Betrachtung der Ableitung des natürlichen Logarithmus (wie ändert sich $\ln(X)$ wenn sich X ändert?)

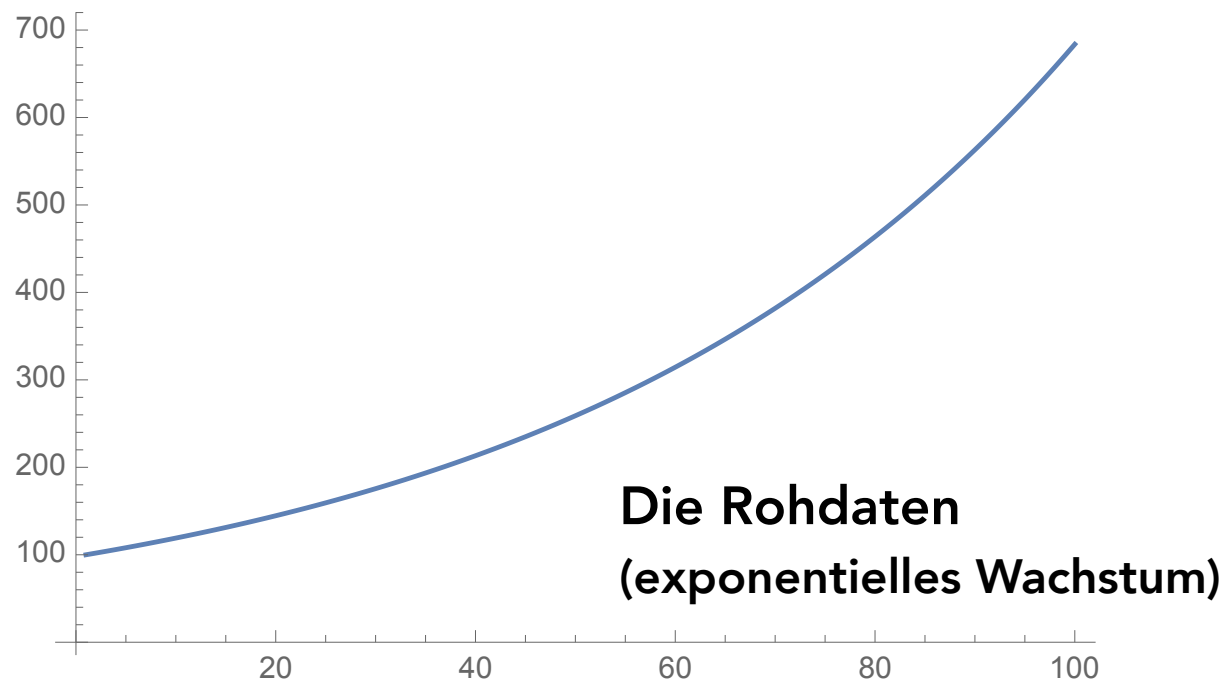
$$\frac{\partial \ln(X)}{\partial X} = \frac{1}{X} \quad \Delta \ln(X) = \Delta X \frac{\partial \ln(X)}{\partial X} = \frac{\Delta X}{X}$$

- Take-Away: Differenz einer logarithmierten Variable misst näherungsweise die relative Änderung der Variable
- Das ist auch für die grafische Darstellung hilfreich: logarithmierte Skalen zeigen relative Änderungen

Zeitreihen und Logarithmus

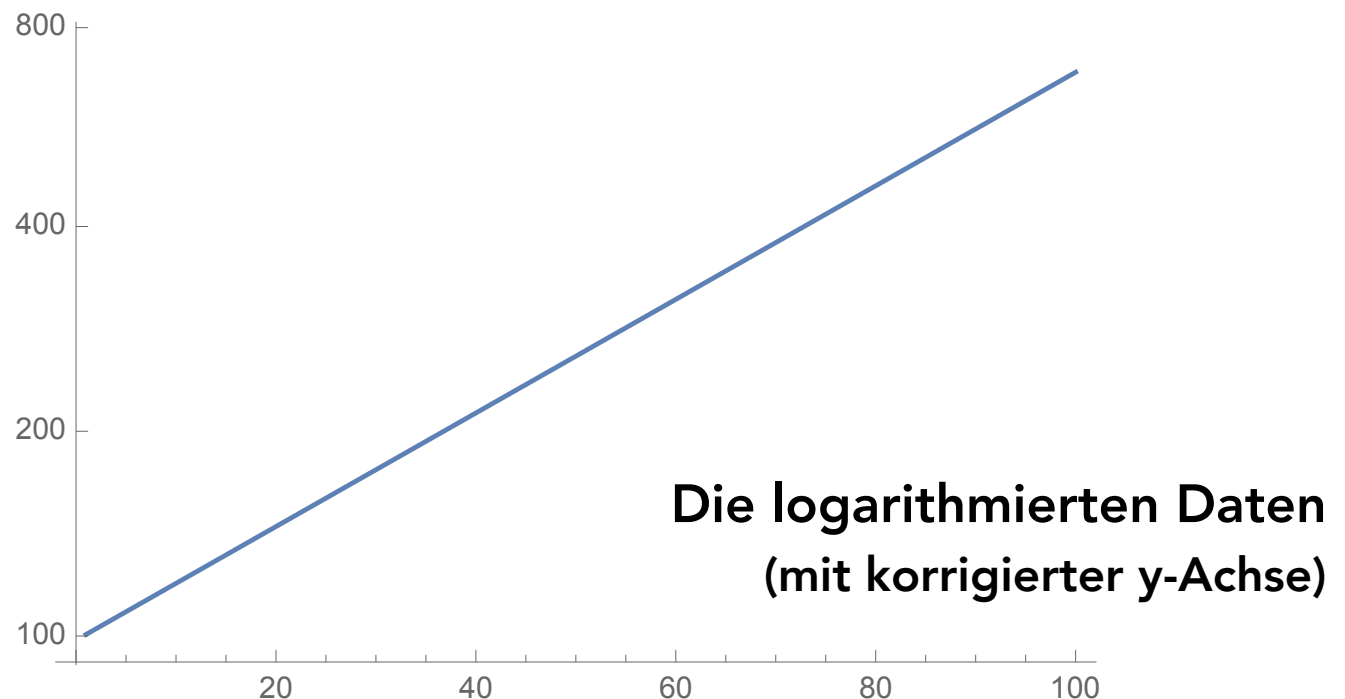
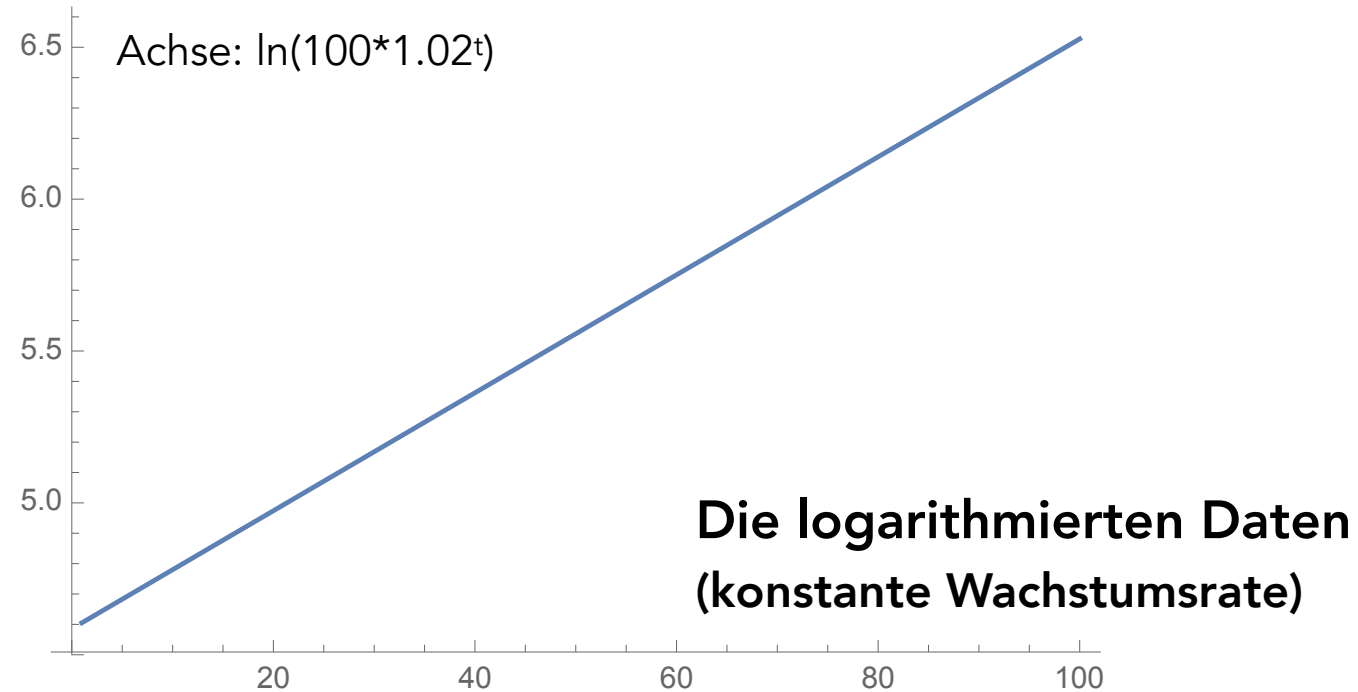
Ein einfaches Beispiel zur Illustration

- Eine Größe wächst mit 2% pro Periode für 100 Perioden (Startwert = 100).



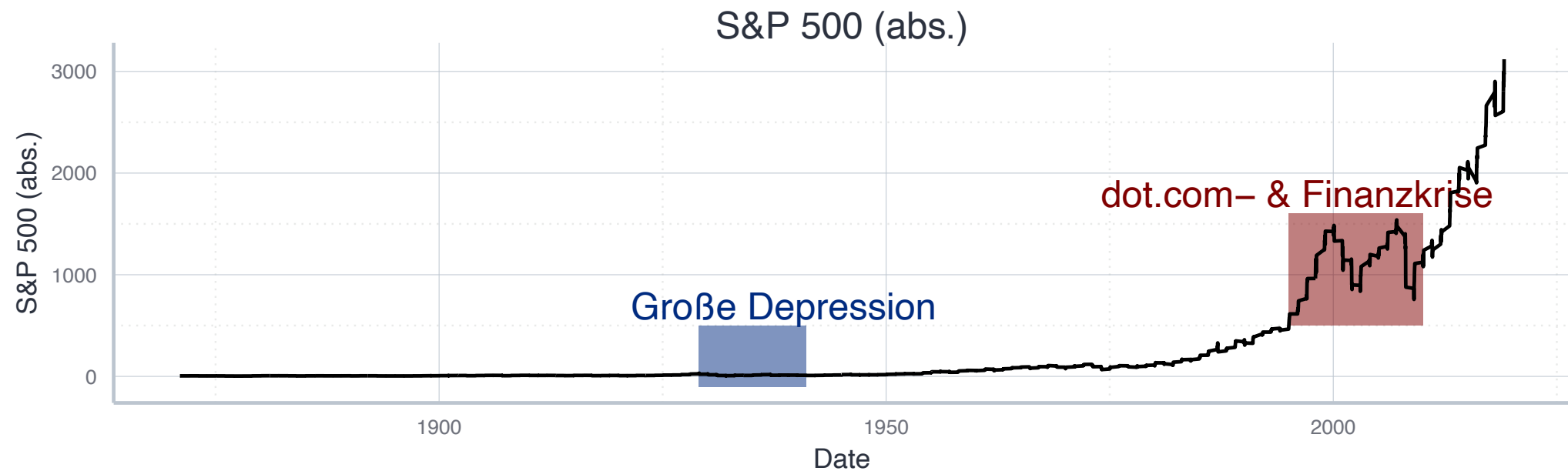
Ergo: Die Steigung der Kurve der logarithmierten Werte zeigt genau die relative Änderung an! (rechts oben)

In absoluten Zahlen bedeutet dies das Wachstum jede Runde ansteigt („Zinseszins“; links oben & rechts unten gut sichtbar)

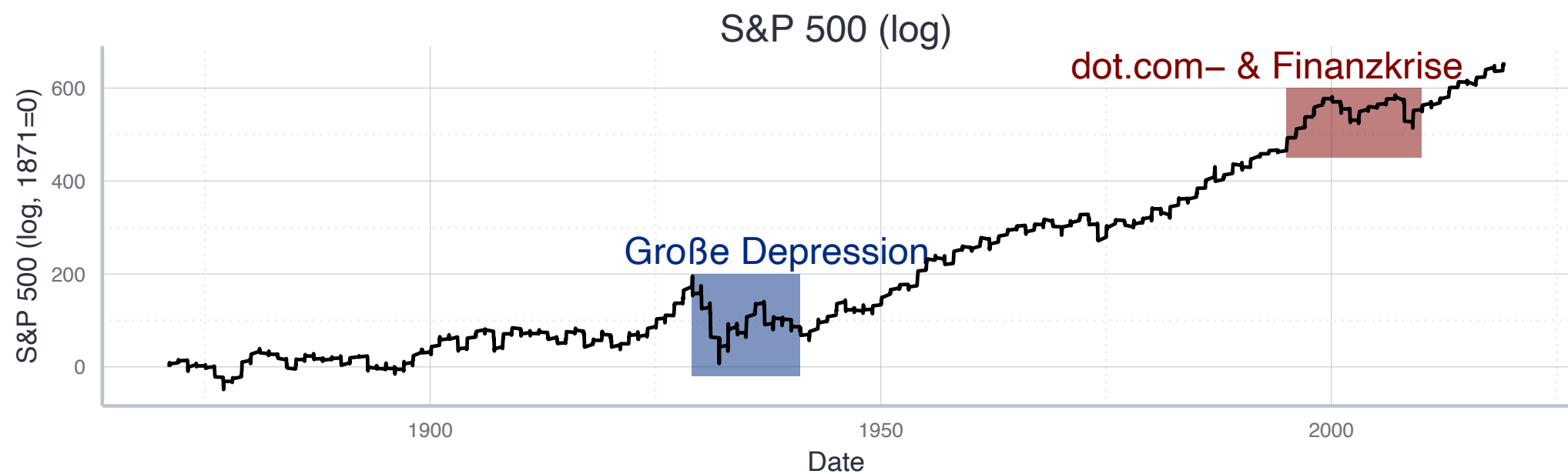


Zeitreihen und Logarithmus

Beispiel: S&P 500 - Rohdaten logarithmierte Daten



Datenquelle: Shiller (2015).



Datenquelle: Shiller (2015).

Änderungsraten und Verwendung des Log's

- In R berechnen Sie Änderungsraten häufig für Gruppen
- Hier bietet sich dann die Verwendung der Funktionen `lead()`, `lag()` und `group_by()` an:

```
##      country      BIP year
## 1  Austria 37941.04 2018
## 2  Germany 35866.00 2018
## 3  Austria 37140.79 2017
## 4  Germany 35477.89 2017
## 5  Austria 36469.39 2016
## 6  Germany 34858.79 2016
## 7  Austria 36129.03 2015
## 8  Germany 34370.64 2015
## 9  Austria 36123.43 2014
## 10 Germany 34076.90 2014
```



```
beispiel_daten <- beispiel_daten %>%
  arrange(country, year) %>%
  group_by(country) %>%
  mutate(BIP_Wachstum = (BIP-dplyr::lag(BIP))/abs(dplyr::lag(BIP))*100) %>%
  ungroup()
beispiel_daten
```

```
## # A tibble: 10 x 4
##   country      BIP year BIP_Wachstum
##   <chr>      <dbl> <int>      <dbl>
## 1 Austria 36123.  2014         NA
## 2 Austria 36129.  2015        0.0155
## 3 Austria 36469.  2016        0.942
## 4 Austria 37141.  2017        1.84
## 5 Austria 37941.  2018        2.15
## 6 Germany 34077.  2014         NA
## 7 Germany 34371.  2015        0.862
## 8 Germany 34859.  2016        1.42
## 9 Germany 35478.  2017        1.78
## 10 Germany 35866.  2018        1.09
```

Optional: Zur Intuition des Logarithmus

Was ist die Interpretation von „e“ und „ln“?

- Für die Eulersche Zahl e gilt: $\lim_{n \rightarrow \infty} (1 + \frac{r}{n})^n = e^r$,
n...Zahl der Perioden
r... Summe der Wachstumsraten über alle Perioden
- **Beispiel:** 4 Jahre mit 2% Wachstum, bedeutet dass $n = 4$ und $r = 8\%$
- e^r hier entspricht also e^{xt} aus vorangegangener Folie (Summe der Wachstumsraten!)
- **Vorteil:** Näherungsweise Berücksichtigung von Zinseszins-Effekten: $e^{0.02 \cdot 4} = 1.0833$
- Natürlicher Logarithmus als inverse Operation zu $e^r = z$:
 - allgemein: $\ln(z)$ ergibt r — „Mit welcher Zahl r muss man e potenzieren um z zu erhalten?“
 - $X_t / X_0 = e^{xt} \rightarrow \ln(X_t / X_0) = \ln(e^{xt}) = xt \rightarrow \ln(X_t) - \ln(X_0) / t = x$
- **Aber:** Eulersche Zahl beschreibt kontinuierliches Wachstum - trifft unseren Fall nicht ganz.
 - Beispiel: $1.02^4 = 1.0824 \neq 1.0833 = e^{0.02 \cdot 4}$ (Annäherung mit e nur näherungsweise!)
 - *Aus welchen Gründen ist der Logarithmus sonst noch attraktiv?*
 - **Antwort:** Weil die logarithmierten Werte in Graphen gut interpretiert werden können.
 - $\ln(X_{t+1}) - \ln(X_t) = \ln(e^x) = x \rightarrow$ Differenz von 0.01, bedeutet eine 1%ige Änderung.

Rechnen mit Erwartungswerten

Erwartungswerte

- **Erwartungswert:** durchschnittliches Ergebnis eines Zufallsprozesses
 - Beispiel Würfel: $E_{W6}(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \dots + \frac{1}{6} \cdot 6 = \frac{1}{6} \sum_{i=1}^n X_i = 3.5$
 - Für eine Stichprobe entspricht der Erwartungswert dem arithmetischen Mittel.
- Wichtige Rechenregeln für Erwartungswerte
 - Manchmal tauchen Erwartungswerte in der statistischen Literatur bzw. manchmal ist es sinnvoll eine Größe als Erwartungswert anzuschreiben um genauer zu sehen was passiert.
 - **Konstante Faktoren:** $E(a + bX) = a + bE(X)$
 - **Summe und Differenz:** $E(X \pm Y) = E(X) \pm E(Y)$
 - **Produkte:** $E(XY) = E(X) \cdot E(Y) + Cov(X, Y)$, d.h. wenn X und Y unabhängig / unkorreliert sind (also $Cov(X, Y) = 0$), kann man die Multiplikation auch aufsplitten, sonst ist das Ganze nicht so trivial.

Beispiele für das Rechnen mit Erwartungswerten

- Kovarianz als Erwartungswert

$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = \sigma_{X,Y}$$

- $\text{Cov}(X, Y) = E((X - E(X)) \cdot (Y - E(Y))) \Leftrightarrow$
- $E(XY) - E(X \cdot E(Y)) - E(E(X) \cdot Y) + E(E(Y) \cdot E(X)) \Leftrightarrow$
- $E(XY) - E(Y) \cdot E(X) - E(X) \cdot E(Y) + E(Y) \cdot E(X) \Leftrightarrow$
- $E(XY) - E(X) \cdot E(Y)$

Wenn X und Y unabhängig sind, dann gilt $E(XY) = E(X) \cdot E(Y)$ und somit $\text{Cov}(X, Y) = 0$.

- Varianz-Kovarianz Matrix (Beschreibung der Fehler im Regressionsverfahren)

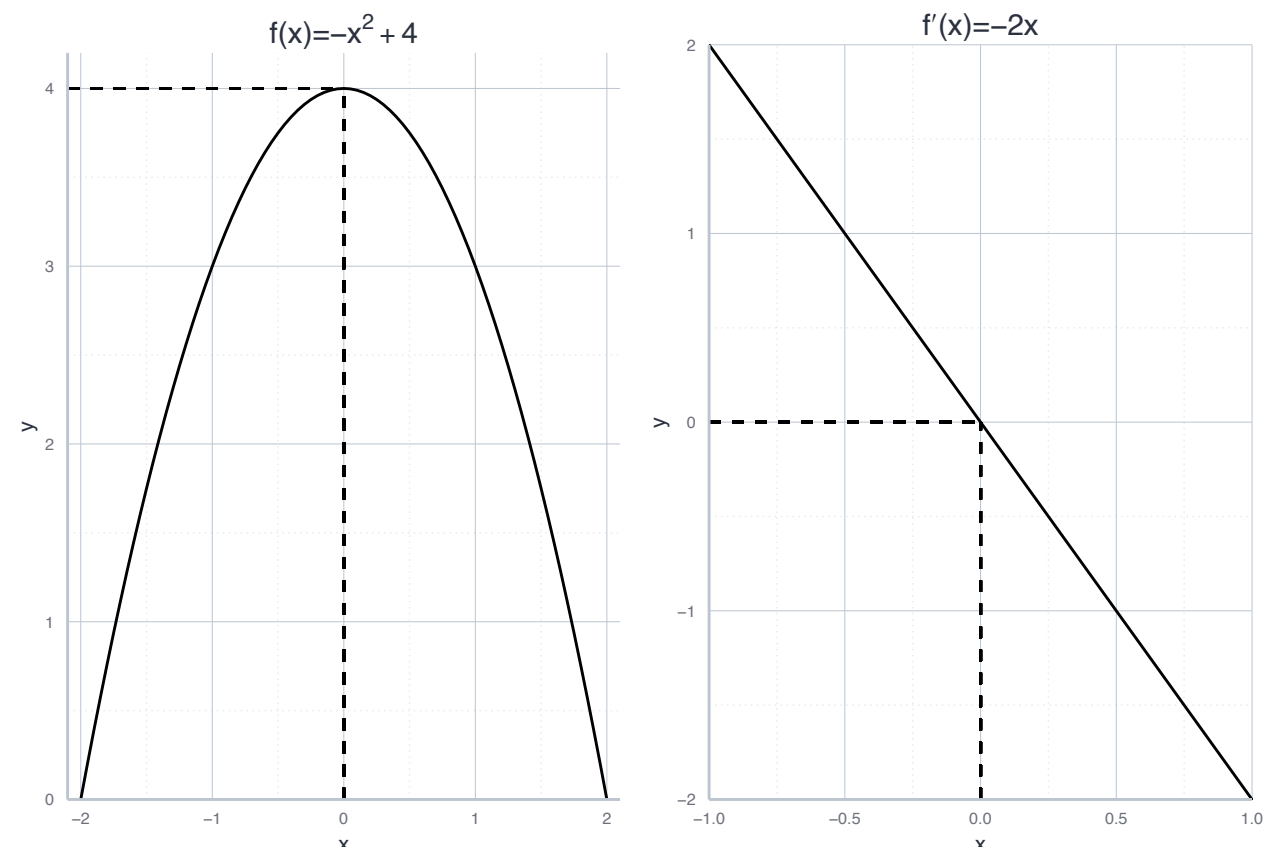
$$E(\epsilon\epsilon' | X) = \begin{pmatrix} \text{Var}(\epsilon_1 | X) & \text{Cov}(\epsilon_1, \epsilon_2 | X) & \text{Cov}(\epsilon_1, \epsilon_3 | X) \\ \text{Cov}(\epsilon_2, \epsilon_1 | X) & \text{Var}(\epsilon_2 | X) & \text{Cov}(\epsilon_2, \epsilon_3 | X) \\ \text{Cov}(\epsilon_3, \epsilon_1 | X) & \text{Cov}(\epsilon_3, \epsilon_2 | X) & \text{Var}(\epsilon_3 | X) \end{pmatrix} \stackrel{? \longrightarrow E(\epsilon_i) = 0}{=} \begin{pmatrix} \epsilon_1^2 & \epsilon_1\epsilon_2 & \epsilon_1\epsilon_3 \\ \epsilon_2\epsilon_1 & \epsilon_2^2 & \epsilon_2\epsilon_3 \\ \epsilon_3\epsilon_1 & \epsilon_3\epsilon_2 & \epsilon_3^2 \end{pmatrix}$$

Differentialrechnung

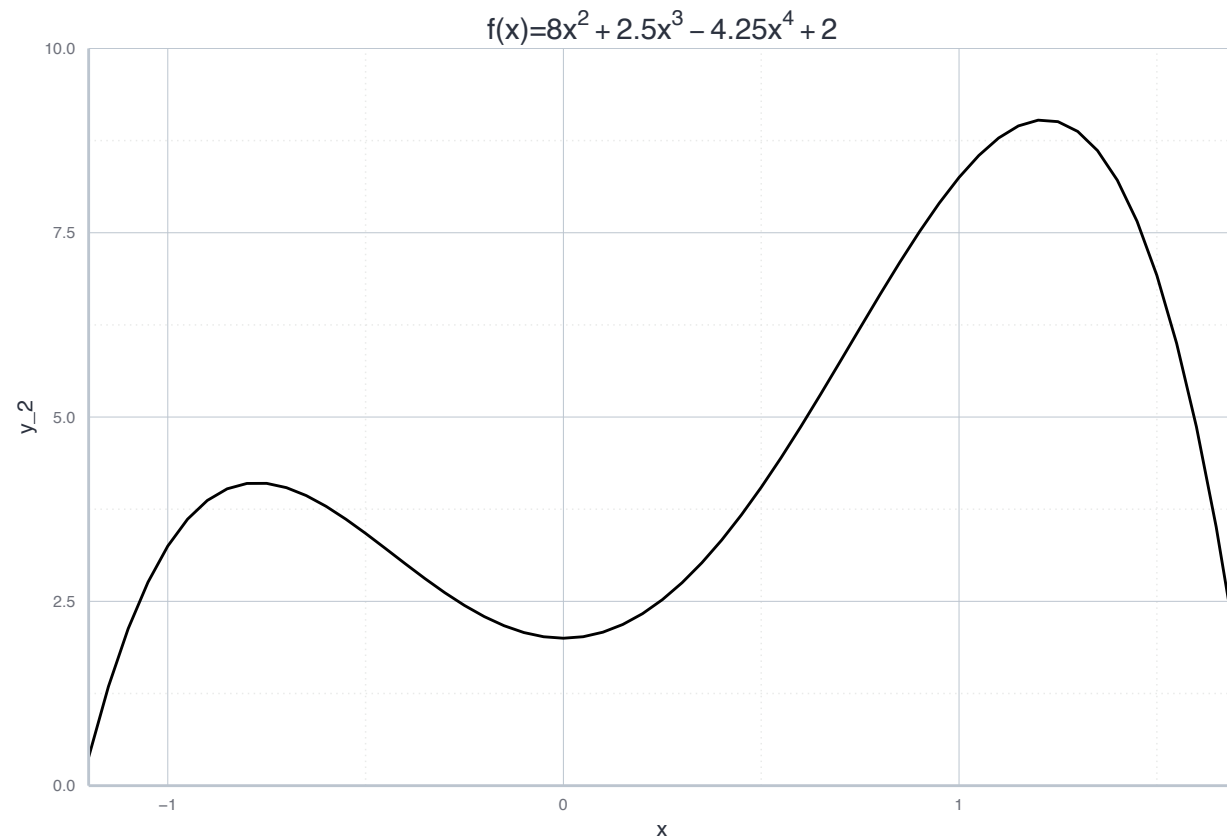
Differentialrechnung

- Differentialrechnung = Analyse lokaler Änderungen von Funktionen
 - Insbesondere: Steigung
 - Hauptanwendungsgebiet: Optimierung
 - Standardprozedur: Ableiten von Funktionen
- Die Ableitung $f'(x)$ einer Funktion $f(x)$ gibt Informationen über die Steigung von $f(x)$ am Punkt x
- Warum ist das für die Optimierung relevant?
 - An Extrempunkten ist die Steigung 0
 - Die **Suche nach Extremwerten von $f(x)$** kann übersetzt werden in die **Suche nach Nullstellen von $f'(x)$**

$$f'(x) \Leftrightarrow \frac{df(x)}{dx} \Leftrightarrow \frac{\partial f(x)}{\partial x} \Leftrightarrow \frac{\partial y}{\partial x}$$



Differentialrechnung



- Nullstellen zu finden ist deutlich einfacher als Extremwerte zu suchen
- Wie leitet man ab?
 - Analytisch: siehe Ableitungsregeln im Skript/Lehrbuch
 - In R: mit der Funktion D()

Differentialrechnung

- Nehmen wir an wir wollen folgende Funktion ableiten:
- $f(x) = x^2 + 3x$
- Wir übersetzen die Funktion zunächst in eine `expression`:

```
f <- expression(x^2+3*x)
f
```

```
x <- 1:5
eval(f)
```

```
#> expression(x^2 + 3 * x)
```

```
#> [1] 4 10 18 28 40
```

- Dann können wir die Funktion mit `D()` ableiten
 - Erstes Argument: die abzuleitende expression
 - Zweites Argument: die Variable nach der wir ableiten wollen

```
D(f, "x")
```

```
#> 2 * x + 3
```

$$\longleftrightarrow \frac{\partial f(x)}{\partial x} = 2x + 3$$

Differentialrechnung

- Sie können die Aufrufe von `D()` auch verschachteln um höhere Ableitungen zu bekommen:

```
D(f, "x")
```

```
#> 2 * x + 3
```



$$\frac{\partial f(x)}{\partial x} = 2x + 3$$

```
D(D(f, "x"), "x")
```

```
#> [1] 2
```



$$\frac{\partial^2 f(x)}{\partial^2 x} = 2$$

- Dies ist hilfreich um genauere Informationen über die Art des zu suchenden Extremwertes zu bekommen.

- Zahl der Extremwerte hängt von Grad des Polynoms ab (höchster Exponent – 1)

1. Ableitung	2. Ableitung	Ergebnis
$f'(x) = 0$	$f''(x) > 0$	Minimum
$f'(x) = 0$	$f''(x) < 0$	Maximum
$f'(x) = 0$	$f''(x) = 0$	Wendepunkt

Optimierungsprobleme in R numerisch lösen

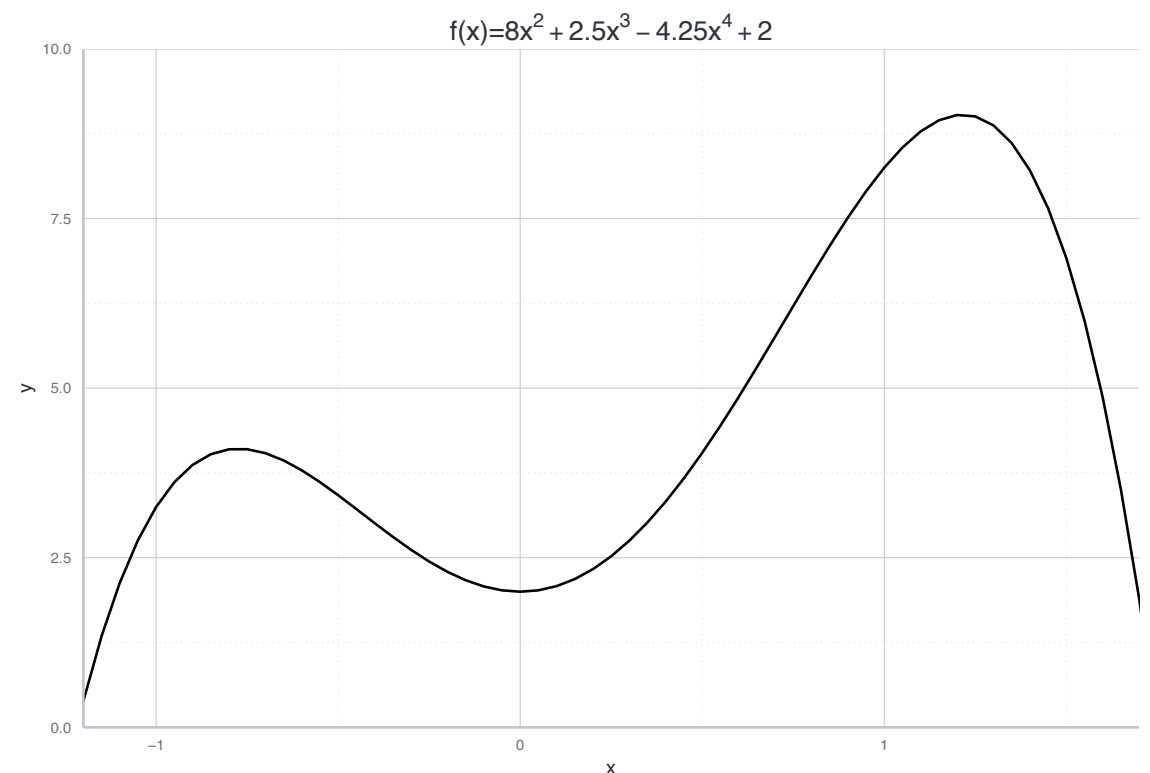
- Wir können in R Optimierungsprobleme auch numerisch lösen
 - *Algorithmische/numerische* Optimierungen gehen mit möglichen Problemen einher.
 - Iterative Suche nach Optima anstatt analytischer Bestimmung.
 - Wir können uns über das Ergebnis nie vollständig sicher sein.
 - Wenn wir den höchsten Punkt der Erde suchen und am Campus Duisburg starten könnte es passieren, dass wir den Kaiserberg als „höchsten“ Punkt finden.

- Betrachten wir folgende Funktion:

$$f(x) = 8x^2 + 2.5x^3 - 4.25x^4 + 2$$

- Für die Optimierung übersetzen wir sie in eine **Funktion**

```
f_1 <- function(x) 8*x^2 + 2.5*x**3 - 4.25*x**4 + 2
```



Optimierungsprobleme in R

$$f(x) = 8x^2 + 2.5x^3 - 4.25x^4 + 2$$

- Für die Optimierung übersetzen wir sie in eine **Funktion**

```
f_1 <- function(x) 8*x^2 + 2.5*x**3 - 4.25*x**4 + 2
```

- Je nach Art der Funktion verwenden wir die passende Optimierungsfunktion:

```
opt_obj <- optimize(f = f_1,  
  lower = -1.25, upper = 1.75,  
  maximum = FALSE)
```

opt_obj

```
## $minimum  
## [1] -7.54766e-06  
##  
## $objective  
## [1] 2
```

TRUE

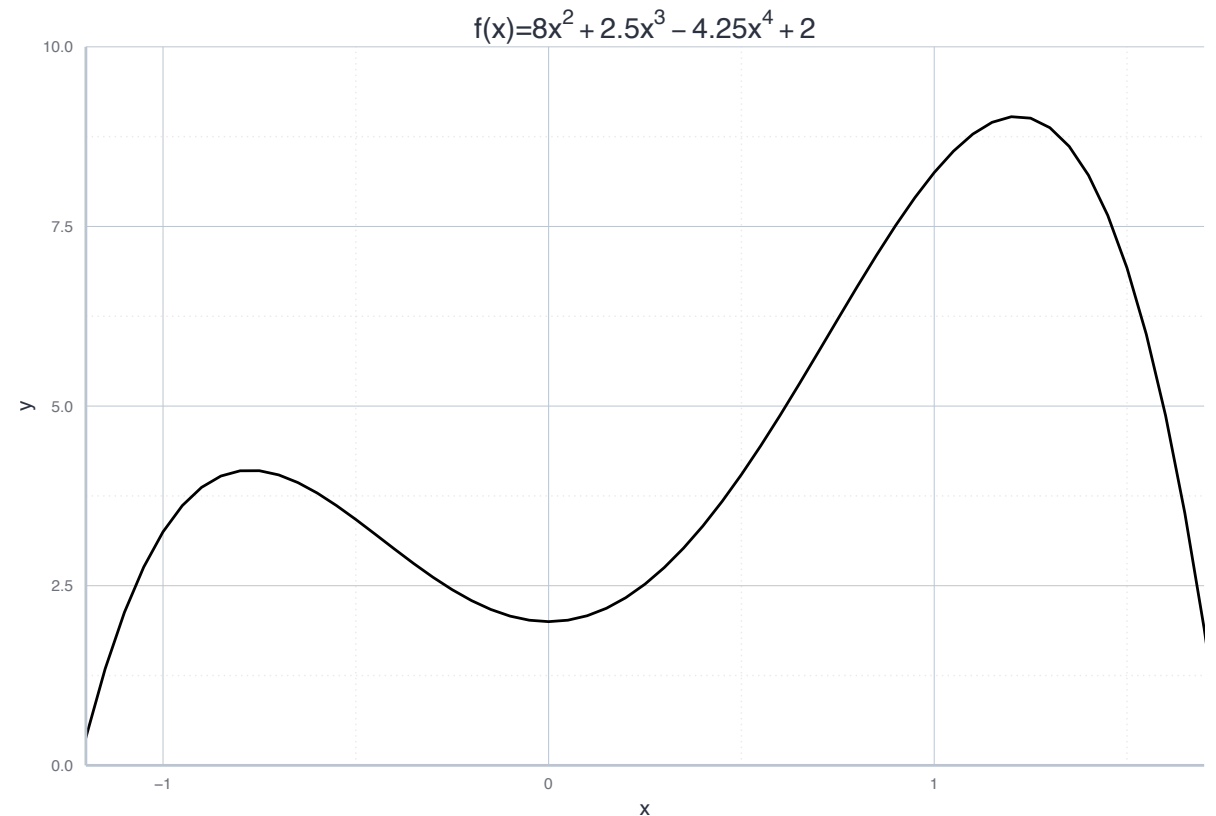
```
## $maximum  
## [1] 1.215492  
##  
## $objective  
## [1] 9.032067
```

Suchbereich
macht den
Unterschied

```
opt_obj <- optimize(f = f_1,  
  lower = -1.25, upper = 0,  
  maximum = TRUE)
```

opt_obj

```
## $maximum  
## [1] -0.7743199  
##  
## $objective  
## [1] 4.108106
```



Partielle und totale Ableitung

- **Partielle Ableitung:** Wie wirkt sich die Änderung einer Variablen aus?
- **Totale Ableitung:** Wie wirkt eine Änderung aller Argumente der Funktion?

- Beispiel: Eine Ertragsfunktion einer Firma

$$R = R(q_1, q_2) = p_1 q_1 + p_2 q_2$$

- **Partielle Ableitungen**

$$\frac{\partial R}{\partial q_1} = p_1 \quad \frac{\partial R}{\partial q_2} = p_2$$

- **Totales Differential**

$$\Delta R = \frac{\partial R}{\partial q_1} \cdot \Delta q_1 + \frac{\partial R}{\partial q_2} \cdot \Delta q_2 = p_1 \Delta q_1 + p_2 \Delta q_2$$

Differentialrechnung - Zusammenfassung

- Optimierungsprobleme tauchen immer wieder auf
 - Einfachere Probleme mit „schöner“ funktionaler Form können wir immer auch analytisch lösen.
 - Kniffligere Optimierungsprobleme kann man in R auch numerisch lösen.
 - Numerische Optimierung ist aber nie vollständig sicher.
 - Das totale Differential kann helfen, wenn eine Variable viele Einflussfaktoren hat.
- Hinweis zur analytische Optimierung
 - Werfen Sie einen Blick in das Skript um einen Überblick über die verschiedenen Optimierungsfunktionen und ihr Anwendungsgebiet zu bekommen.
- Hinweis zur numerischen Optimierung
 - Hier ist oft der Suchbereich entscheidend für das Ergebnis.

Lineare Algebra

Lineare Algebra

- In der linearen Algebra studiert man vor allem lineare Funktionen
 - Eine Gruppe ähnlicher linearer Funktionen kann in Matrizenform ausgedrückt werden
- Die Rechenregeln für Matrizen werden als **Matrizenalgebra** bezeichnet
 - Kann in R einfach durchgeführt werden (siehe Skript)
- Wir werden die Grundlagen der Matrizenalgebra hier nicht extensiv wiederholen (können)
 - Aber: Einige Grundgedanken und zwei Beispiele
 - Ein einfaches keynesianisches Modell
 - Einfaches Input-Output-Modell (Leontief!)
 - Der OLS Schätzer
 - Hilfreich, da alle fortgeschrittenen Ökonometrietexte in Matrizenform geschrieben sind

Was ist und was kann eine Matrix?

Grundsätzliche Aspekte

- Man kann sich eine Matrix als zwei-dimensionalen Vektor denken

$$\mathbf{v}' = (v_1, v_2, \dots, v_n)$$
$$\mathbf{A}^{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Eine Matrix hat „Dimensionen“ (m Zeilen und n Spalten)

- Man kann mit Matrizen rechnen
 - Summe** und **Differenz** ergeben sich aus der Addition / Subtraktion der Elemente (Dimensionen müssen zusammenpassen!)
 - Matrixmultiplikation** geht auch: Es gilt „Zeile mal Spalte“ (Reihenfolge und Transponieren spielen eine Rolle! Matrizen müssen „zusammenpassen“!)
 - Division** gibt es auch: Man braucht dafür aber eine „Inverse“ A^{-1}
Die Inverse existiert aber nur, wenn die Determinante $|A| \neq 0$;-)

Was ist und was kann eine Matrix?

Basics zur Matrixmultiplikation

- Matrixmultiplikation: (jede) Zeile mal (jede) Spalte
 - **Erste Konsequenz:** Reihenfolge und Transponierung spielen eine Rolle → **z.B. Residuen**

Transponierter Spaltenvektor
= Zeilenvektor der Residuen

$$\underbrace{e'}_{\text{Spaltenvektor der Residuen}} \cdot \underbrace{e}_{\text{Zeilenvektor der Residuen}} = (e_1, e_2, \dots, e_n) \cdot \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = e_1^2 + e_2^2 + \dots + e_n^2 = RSS$$

Ein Skalar (1x1 Matrix) als Ergebnis:
Summe der quadrierten Residuen

$$e \cdot e' = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \cdot (e_1, e_2, \dots, e_n) = \begin{pmatrix} e_1^2 & e_1 e_2 & \dots & e_1 e_n \\ e_2 e_1 & e_2^2 & \dots & e_2 e_n \\ \vdots & \vdots & \ddots & \vdots \\ e_n e_1 & e_n e_2 & \dots & e_n^2 \end{pmatrix}$$

Umgekehrte Reihung:
n x n Matrix als Ergebnis:
„Varianz-Kovarianz Matrix“

- **Zweite Konsequenz:** Matrizen müssen „zusammenpassen“ – z.B. $A^{m \times n} \cdot B^{p \times q}$
- Damit Multiplikation klappt muss $n = p$ gelten (Spalten von A = Zeilen von B).
- Die resultierende Matrix $C^{m \times q}$ hat dann die Dimensionen $m \times q$.

Was ist und was kann eine Matrix?

Gleichungssysteme im Matrixform und die Inverse (Division)

- Stellen wir uns ein einfaches **lineares Gleichungssystem** vor, z.B

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= d_1 \\ a_{21}x_1 + a_{22}x_2 &= d_2 \end{aligned} \Leftrightarrow \underbrace{\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\text{Gleichungssystem in Matrixform}} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \Leftrightarrow \mathbf{A}^{2 \times 2} \cdot \mathbf{x} = \mathbf{d} \Leftrightarrow \boxed{\mathbf{Ax} = \mathbf{d}}$$

Kompakte Darstellung
eines Gleichungssystems!

- Davon ausgehend kann man Gleichungssysteme sehr schnell „lösen“

$$\begin{aligned} \mathbf{Ax} = \mathbf{d} &\Leftrightarrow \mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{d} \Leftrightarrow \underbrace{\mathbf{Ix}}_{\text{Identitätsmatrix (analog zu 1)}} = \mathbf{A}^{-1}\mathbf{d} \Leftrightarrow \\ \mathbf{x} &= \mathbf{A}^{-1}\mathbf{d} \end{aligned} \quad \Rightarrow \quad \mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

- **Gleichungssysteme:** Drei mögliche Outcomes

- (1) Eindeutige Lösung $\rightarrow$ Inverse $\mathbf{A}^{-1}$ existiert und Determinante $|\mathbf{A}| \neq 0$ („nicht-singulär“).
- (2) Keine Lösung & (3) unendlich viele Lösungen $\rightarrow$ Inverse $\mathbf{A}^{-1}$ existiert nicht und Determinante $|\mathbf{A}| = 0$ („singuläre Matrix“).

Bezug zu OLS-Annahme: Keine perfekte Multikollinearität!

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 1: Das einfache keynesianische Modell ;-)

- Wir können das einfache keynesianische Modell mit Staat als ein lineares Gleichungsmodell mit drei endogen bestimmten Variablen („Unbekannten“) schreiben:

$$Y = C + I + G \quad \text{Aggregierte Nachfrage}$$

$$C = a + b(Y - T) \quad \text{Konsumgleichung}$$

$$T = \tau Y \quad \text{Einkommenssteuer}$$

- In Matrixform:

$$Y - C = I + G \Leftrightarrow Y = C + I + G$$

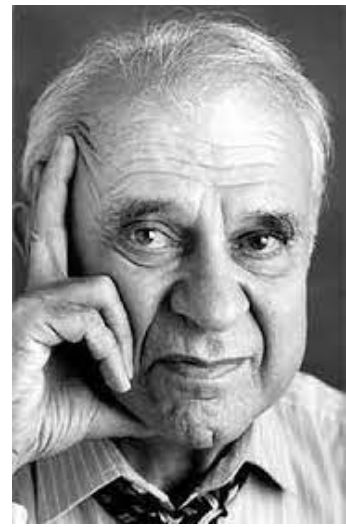


Zeilenweise
Ausmultiplizieren!

$$\begin{pmatrix} 1 & -1 & 0 \\ -b & 1 & b \\ -\tau & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} Y \\ C \\ T \end{pmatrix} = \begin{pmatrix} I + G \\ a \\ 0 \end{pmatrix}$$

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 2: Input-Output Tabellen und die Leontief-Inverse



- Stellen wir uns eine Ökonomie mit drei Sektoren vor:

$$x_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + d_1 \longrightarrow x_1(1 - a_{11}) - a_{12}x_2 - a_{13}x_3 = d_1$$

$$x_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + d_2 \quad -a_{21}x_1 + x_2(1 - a_{22}) - a_{23}x_3 = d_2$$

$$x_3 = a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + d_3 \quad -a_{31}x_1 - a_{32}x_2 + x_3(1 - a_{33}) = d_3$$

Gut (Input) Technischer Koeffizient: Wieviel x_3 für ein x_2 ? Gut (Output) Endnachfrage x_3

- In Matrixform: $A^* \cdot x = d \Leftrightarrow \begin{pmatrix} (1 - a_{11}) & -a_{12} & -a_{13} \\ -a_{21} & (1 - a_{22}) & -a_{23} \\ -a_{31} & -a_{32} & (1 - a_{33}) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$

Welche Gesamtproduktion ergibt sich bei gegebener Nachfrage nach Endprodukten?

Wenn $A^{n \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$, dann $A^* = (I - A) \Rightarrow (I - A) \cdot x = d \Leftrightarrow$

Leontief-Inverse! $(I - A)^{-1} \cdot d = x$

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer

- Wir kennen die lineare Regression bislang in der Gleichungsform:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- Für beliebig viele erklärende Variablen (hier: p) wäre das:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \epsilon$$

- Wenn wir jede einzelne Beobachtung y_i ausschreiben ergibt sich:

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p} + \epsilon_1, \quad i = 1, \dots, n$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p} + \epsilon_2, \quad i = 1, \dots, n$$

$$\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots$$

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \epsilon_n, \quad i = 1, \dots, n$$

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_p x_{1p} + \epsilon_1, & i &= 1, \dots, n \\ y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_p x_{2p} + \epsilon_2, & i &= 1, \dots, n \\ &\vdots & & \vdots \\ y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_p x_{np} + \epsilon_n, & i &= 1, \dots, n \end{aligned}$$

Juhu, das kann man als Matrix anschreiben!
Was für ein Fest ;-)

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

- Oder ganz kompakt: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer

- Betrachten wir ein Beispiel mit folgendem Datensatz:

##	Auto	Verbrauch	PS	Zylinder
## 1:	Ford Pantera L	15.8	264	8
## 2:	Ferrari Dino	19.7	175	6
## 3:	Maserati Bora	15.0	335	8
## 4:	Volvo 142E	21.4	109	4

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \epsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \epsilon_2$$

$$y_3 = \beta_0 + \beta_1 x_{31} + \beta_2 x_{32} + \epsilon_3$$

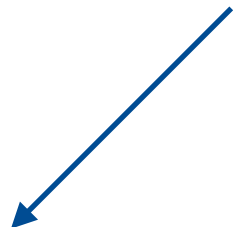
$$y_4 = \beta_0 + \beta_1 x_{41} + \beta_2 x_{42} + \epsilon_4$$

$$15.8 = \beta_0 + \beta_1 264 + \beta_2 8 + \epsilon_1$$

$$19.7 = \beta_0 + \beta_1 175 + \beta_2 6 + \epsilon_2$$

$$15 = \beta_0 + \beta_1 335 + \beta_2 8 + \epsilon_3$$

$$21.4 = \beta_0 + \beta_1 109 + \beta_2 4 + \epsilon_4$$

$$\begin{pmatrix} 1 & 264 & 8 \\ 1 & 175 & 6 \\ 1 & 335 & 8 \\ 1 & 109 & 4 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix} = \begin{pmatrix} 15.8 \\ 19.7 \\ 15.0 \\ 21.4 \end{pmatrix}$$


Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer

$$\begin{pmatrix} 1 & 264 & 8 \\ 1 & 175 & 6 \\ 1 & 335 & 8 \\ 1 & 109 & 4 \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{pmatrix} = \begin{pmatrix} 15.8 \\ 19.7 \\ 15.0 \\ 21.4 \end{pmatrix}$$

- Die Lösung für das OLS-Minimierungsproblem in Matrixschreibweise ist

$$(X'X)^{-1} (X'y) = \hat{\beta} \rightarrow \text{klassische Formulierung des OLS Schätzers}$$

- Wir können das in R händisch lösen:

```
ols_X <- matrix(c(1, 264, 8, 1, 175, 6, 1, 335, 8, 1, 109, 4),
               ncol = 3, byrow = T)
ols_y <- matrix(c(15.8, 19.7, 15.0, 21.4), ncol = 1)

solve(t(ols_X) %*% ols_X) %*% t(ols_X) %*% ols_y

##           [,1]
## [1,] 26.37086491
## [2,] -0.01783627
## [3,] -0.68592421
```

- Oder die Funktion `lm()` benutzen:

```
lm(Verbrauch~PS+Zylinder, data = ols_beispiel)

##
## Call:
## lm(formula = Verbrauch ~ PS + Zylinder, data = ols_beispiel)
##
## Coefficients:
## (Intercept)          PS          Zylinder
##    26.37086    -0.01784    -0.68592
```

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer (Optional: Herleitung)

- Wie kommen wir von $\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\epsilon}$ zu $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} (\mathbf{X}'\mathbf{y})$?
- Die Residuen sind gegeben durch: $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$
- Wir können die Summe der quadrierten Residuen (RSS) schreiben als:

$$\mathbf{e}' \cdot \mathbf{e} = (e_1, e_2, \dots, e_n) \cdot \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = e_1^2 + e_2^2 + \dots + e_n^2 = RSS$$

- Dann kombinieren wir beide Ausdrücke:

$$\begin{aligned} \mathbf{e}'\mathbf{e} &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\ &= \mathbf{y}'\mathbf{y} - 2\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned}$$

- Da wir die Summe der Residuen minimieren wollen leiten wir diesen Ausdruck ab

Cooler Beispiele für die Anwendung von Matrizen

Beispiel 3: Regression und OLS-Schätzer (Optional: Herleitung)

$$\begin{aligned} e'e &= (Y - X\hat{\beta})' (Y - X\hat{\beta}) = y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\ &= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta} \end{aligned}$$

- Da wir die Summe der Residuen minimieren wollen, leiten wir diesen Ausdruck ab

$$\frac{\partial e'e}{\partial \hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

- Das wiederum können wir umformen:

$$2X'X\hat{\beta} = 2X'y$$

$$X'X\hat{\beta} = X'y$$

- Da $(X'X)^{-1} (X'X) = I$ ergibt sich hieraus dann der OLS-Schätzer:

$$(X'X)^{-1} X'X\hat{\beta} = (X'X)^{-1} X'y$$

$$\hat{\beta} = (X'X)^{-1} (X'y)$$

Formale Verfahren: Resümee

- Formale Methode als Grundlage vieler methodischer Techniken und Überlegungen
 - Relative Änderung, geometrisches Wachstum, multiplikative Dynamik und der Logarithmus.
 - Matrizenmathe als Tool für ökonomische Theorie und Ökonometrie (insbesondere auch die Übersetzung zwischen beiden Felder).
 - Erwartungswerte und Ihre Rolle in der Statistik.
 - Ableitungen als Kernelement der Neoklassik ;-)
- Eine Software kann helfen
 - z.B. um abstrakte Funktionen zu visualisieren
 - z.B. um mal fröhlich ein numerisches Beispiel durchzurechnen.
 - Die eigentlichen Berechnungen sind oft einfach mit *R* durchzuführen.

Wiederholungsfragen zum Selbststudium

Wiederholungsfragen

- Was ist der Unterschied zwischen arithmetischem und geometrischem Mittel?
- Wie kann die Differenz zweier logarithmierter Werte interpretiert werden?
- Geben Sie ein Beispiel wann die Darstellung einer Größe in logarithmierter Form besonders nützlich erscheint.
- Was ist der Erwartungswert (a) eines 20 seitiges Würfels und (b) eines Wurfs mit zwei sechsseitigen Würfeln?

Wiederholungsfragen

- Wie hängen Optimierungsprobleme und Differenzialrechnung miteinander zusammen?
- Welche zwei Schritte sind in R notwendig um eine Funktion abzuleiten?
- Welche zwei Schritte sind in R notwendig um eine Funktion zu optimieren?
- Was verstehen wir unter einer partiellen Ableitung, was unter einem totalen Differential?
- Welche Gefahr besteht immer wenn wir ein Optimierungsproblem nicht analytisch sondern nur numerisch lösen?

Wiederholungsfragen

- Schreiben Sie die folgenden beiden Gleichungen in Matrixform auf:
 1. $Y = C + I + G$
 2. $C = a + bY$
- Was verstehen wir allgemein unter der 'Lösung' eines linearen Gleichungssystems?
- In welche allgemeine Form können Sie ein lineares Gleichungssystem umformen um die Lösung zu erhalten?
- Welche möglichen Arten von Lösungen gibt es für ein lineares Gleichungssystem (und was hat das mit der Inverse zu tun)?
- Wie lautet die Formulierung der OLS-Schätzgleichung wie wir sie letzte Session kennen gelernt haben und wie sieht die Übersetzung in Matrixschreibweise aus?